

# Transcription fluctuation effects on biochemical oscillations

Ryota Nishino, Takahiro Sakaue, and Hiizu Nakanishi

*Department of Physics, Kyushu University 33, Fukuoka 812-8581, Japan*

---

## Abstract

Some biochemical systems show oscillatory behavior. These systems often consist of negative feedback loops with repressive transcription regulation. Such biochemical systems have distinctive characteristics in comparison with ordinary oscillatory chemical systems: i) the numbers of molecules involved are small, ii) there are typically only a couple of genes in a cell with a finite regulation time scale. Due to the fluctuations caused by these features, the system behavior can be quite different from the one obtained by deterministic rate equations, because the rate equations ignore molecular fluctuations and thus are exact only in the infinite molecular number limit. The molecular fluctuations on a free-running circadian system have been studied by Gonze et al. (2002) by introducing a scale parameter  $\Omega$  for the system size. They consider, however, only the first effect, assuming that the gene process is fast enough for the second effect to be ignored, but this has not been examined systematically yet. In this work, we study fluctuation effects due to the finite gene regulation time by introducing a new scale parameter  $\tau$ , which we take as the unbinding time of a nuclear protein from the gene. We particularly focus on the case where the fluctuations due to small molecular numbers can be ignored. In simulations on the same system studied by Gonze et al., we find the system is unexpectedly sensitive to the fluctuation in the transcription regulation; the period of oscillation fluctuates about 30 min even when the regulation time scale  $\tau$  is around 30 s, that is even smaller than 1/1000 of its circadian period. We also demonstrate that the distribution width for the oscillation period and the amplitude scales with  $\sqrt{\tau}$ , and the correlation time of the oscillation scales with  $1/\tau$  in the small  $\tau$  regime. The relative fluctuations for the period are about half of that for the amplitude, namely, the periodicity is more stable than the amplitude.

*Keywords:* circadian rhythm, molecular fluctuation, model simulation

---

## Introduction

One of the outstanding features in biological systems is that the systems often operate on surprisingly small numbers of active molecules, yet they seem to work quite reliably. This is especially intriguing in the case where the chemical reaction system involves a gene transcription because there are typically only a couple of genes in a cell.

One example is a circadian system, which shows a rhythmic behavior of approximately 24-hour periodicity. It is a universal feature of biological systems and known to be very accurate and robust against external and internal perturbations (Dunlap, 1999; Young, 2000). Its biochemical mechanisms have been proposed in several systems (Goldbeter, 1995, 1996; Leloup et al., 1999), and most of them are based on a time-delayed negative feedback loop of a biochemical reaction network which includes transcription regulations. Some of the protein molecules are expected to be very small in number, and the number of each gene is typically of the order of one in a cell and does not scale with the cell size, thus it is surprising that the circadian system is capable of maintaining its extraordinary regularity especially in the case of a single cell organism (Barkai and Leibler, 2000).

The effects of molecular fluctuations on the circadian system has been studied by Gonze et al. (2002b,a) by Monte Carlo simulations using the Gillespie method (Gillespie, 1976, 1977) with the scale parameter  $\Omega$  for the molecular numbers. By simulating the system with various values of the scaling parameter  $\Omega$ , they demonstrated that the system shows reasonably coherent oscillation as long as the system contains more than several tens of mRNA, thus concluded that their system are fairly robust against molecular fluctuations.

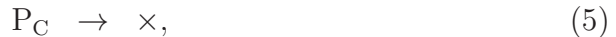
They examined the system rather systematically based upon a standard method to study the stochastic nature of chemical reactions (Nicolis, 1972) by scaling the reaction rates in the way to keep the rate equation unchanged. However, there is an ambiguity in the treatment of the gene regulation process because the number of genes should not scale with other protein numbers. They scaled the reaction rates proportional to  $\Omega$  for the gene processes, namely, they made the gene regulation times infinitely fast in the large  $\Omega$  limit, thus the system dynamics reduce to the one described by the corresponding rate equations without the gene processes (Gonze et al., 2002b).

This could be justified only when the gene processes are so fast that they do not cause significant fluctuation on the system behavior. In fact, it is not reasonable to assume that the time scale of the gene regulation depends upon the scale parameter  $\Omega$ , if you think of it as the cell volume, because the time scale with which a regulatory protein binds to the operator site is determined by the protein concentration, and the time scale with which the protein unbinds is determined by its binding energy.

In this work, in order to analyze fluctuations from the two distinct origins separately, we introduce a new scale parameter  $\tau$  in addition to  $\Omega$ . The scale parameter  $\tau$  scales the binding/unbinding time of the gene regulatory protein. Thus, these two parameters,  $\tau$  and  $\Omega$ , control the two distinct fluctuation sources that exist in the biological systems, namely,  $\Omega$  controls the fluctuations due to the finite molecular numbers while  $\tau$  controls the fluctuations due to the finite gene regulation times. We perform the Monte Carlo simulations on the same system as the one studied by Gonze et al. (2002a), focusing on the latter effect, and demonstrate that significant fluctuation can arise from the stochasticity in the gene process alone. We also examine how the fluctuation scales with  $\tau$ .

## Model

The model we study is the simplest version of core model for a circadian system that consists of a gene  $G$ , mRNA  $M$ , cytosolic protein  $P_C$ , and nuclear protein  $P_N$  (Fig.1). The biochemical reactions for these elements are given by



The parameter  $n$  is the number of nuclear proteins  $P_N$  that bind to suppress the gene, i.e. Hill coefficient for the gene activity; we adopt  $n = 4$  for most of the calculation as Gonze et al. (2002b,a).

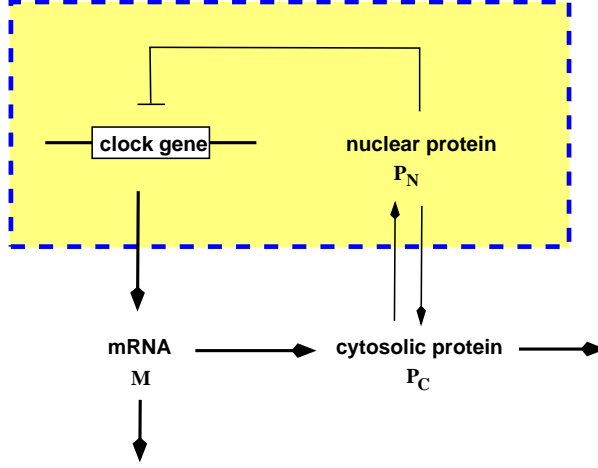


Figure 1: Simplified core model for a circadian system.

Now, we introduce the two scaling parameter  $\Omega$  and  $\tau$ ;  $\Omega$  scales the reaction rates so that the numbers of mRNA and the proteins become proportional to it, and  $\tau$  scales the binding/unbinding time of the nuclear protein to the gene operator site. The transition rates for each reaction are listed in Table 1, where we define the variables  $G$ ,  $M$ ,  $P_C$ , and  $P_N$  to represent the numbers of active genes, mRNA, cytosolic proteins, and the nuclear proteins in a single cell, respectively. The gene variable  $G$  takes either 1 or 0 values, depending upon the active state (G) or the inactive state ( $GP_{Nn}$ ), respectively. Note that we employ Michaelis-Menten enzymatic reactions for the degradation processes. The first two reaction rates in Table 1 are for the gene regulation and proportional to  $1/\tau$ , but do not scale with  $\Omega$  because we assume only one gene in a cell<sup>1</sup>. The ratio of the binding and the unbinding times is determined in the way that the corresponding average behavior described by the rate equation remains the same with the original system in the small  $\tau$  limit. On the other hand, the gene transcription activity in the third reaction is scaled as  $v_s\Omega$  in order that the numbers of mRNA and the proteins should be proportional to  $\Omega$ .

If we naively write down differential equations for the time evolution,

<sup>1</sup> Gonze et al. (2002b) used  $\Omega$  as a scaling parameter in the places where we use  $1/\tau$ .

ignoring the fact that the variables are integers, we would have

$$\tau \frac{dG}{dt} = (1 - G) - \left( \frac{P_N/\Omega}{K_I} \right)^n G, \quad (7)$$

$$\frac{dM}{dt} = v_s \Omega G - v_m \Omega \frac{M}{K_m \Omega + M}, \quad (8)$$

$$\frac{dP_C}{dt} = k_s M - v_d \Omega \frac{P_C}{K_d \Omega + P_C} - k_1 P_C + k_2 P_N, \quad (9)$$

$$\frac{dP_N}{dt} = k_1 P_C - k_2 P_N \quad (10)$$

$$+ \frac{n}{\tau} \left[ (1 - G) - \left( \frac{P_N/\Omega}{K_I} \right)^n G \right]. \quad (11)$$

For ordinary chemical reactions without a gene transcription process, the stochastic dynamics should be well described by such equations for the large  $\Omega$  case, where the numbers of molecules are large. However, in the present system, the number of gene is one and does not scale with  $\Omega$ , thus the stochastic nature remains even in the case of the infinite  $\Omega$  as long as  $\tau$  is finite.

*Large  $\Omega$  limit:*

Supposing the scale parameter  $\Omega$  as a cell volume, we define the “concentrations” of mRNA and the proteins as,

$$[M] \equiv \frac{M}{\Omega}, \quad [P_C] \equiv \frac{P_C}{\Omega}, \quad [P_N] \equiv \frac{P_N}{\Omega}, \quad (12)$$

and write down the rate equations for them as

$$\frac{d[M]}{dt} = v_s G - v_m \frac{[M]}{K_m + [M]}, \quad (13)$$

$$\frac{d[P_C]}{dt} = k_s [M] - v_d \frac{[P_C]}{K_d + [P_C]} - k_1 [P_C] + k_2 [P_N], \quad (14)$$

$$\frac{d[P_N]}{dt} = k_1 [P_C] - k_2 [P_N], \quad (15)$$

where we have ignored the term of the order of  $n/\Omega$  in Eq.(15).

For ordinary chemical reactions, we expect that the deterministic dynamics represented by the rate equations would describe the system accurately in the large  $\Omega$  limit, because the effect of molecular fluctuation becomes negligible. However, for the present case, the system remains stochastic even in the large  $\Omega$  limit because the variable  $G$  remains stochastic.

*Small  $\tau$  limit:*

In the case where  $\tau$  is much smaller than any other time scales in the system, the system reduces to the one studied by Gonze et al. (2002a). This can be seen by introducing the time dependent average value of  $G$ , denoted by  $G_{\text{Av}}(t)$ , i.e. the time average of  $G$  over the longer time scale than  $\tau$  but shorter than other time scales in the system. Its value is given by the condition that the first two reactions in Table 1 are equilibrated,

$$G_{\text{Av}}(t) = \frac{1}{1 + ([\text{P}_{\text{N}}]/K_I)^n}. \quad (16)$$

Then the system dynamics are given by the stochastic dynamics of reaction 1  $\sim$  6 in Table 1 with  $G$  replaced by  $G_{\text{Av}}$ .

If we take the large  $\Omega$  limit on top of this, we obtain the deterministic rate equations given by Eqs.(13)~(15) with  $G$  being replaced by  $G_{\text{Av}}(t)$  of Eq.(16).

## Simulations and results

In order to examine the effect of gene fluctuations, we have performed numerical simulations for various values of  $\tau$  and  $\Omega$ . We examine two cases: (i) the case where both  $\tau$  and  $\Omega$  are finite, and (ii) the case where  $\tau$  is finite but in the large  $\Omega$  limit. In the first case, the fully stochastic dynamics are given by Table 1; for these we employ the Gillespie method (Gillespie, 1976, 1977). In the second case, the concentrations  $[\text{M}]$ ,  $[\text{P}_C]$ , and  $[\text{P}_N]$  follow the deterministic dynamics while the gene process remains stochastic. In this case, we integrate the rate equations (13)~(15) using Runge-Kutta method, but at every time step of the length  $\Delta t$ , the gene variable  $G$  is subject to a trial for change according to the probability  $w\Delta t$  under Poisson process with  $w$  being the transition rate given in the first two processes in Table 1.

Figure 2 shows the system behaviors for various values of  $\Omega$  with  $\tau = 0.01\text{h}$ . The other reaction parameters are the same as those used by Gonze et al. (2002a). The plots in the left column are the time variations of the concentrations of mRNA (solid lines) and the cytosolic protein (dashed lines), and those in the right column show the oscillation trajectories projected on the  $M/\Omega - P_C/\Omega$  plane of the phase space. The fluctuation decreases as  $\Omega$  increases, but it remains finite even in the infinite  $\Omega$  case because of the fluctuation from the gene activity.

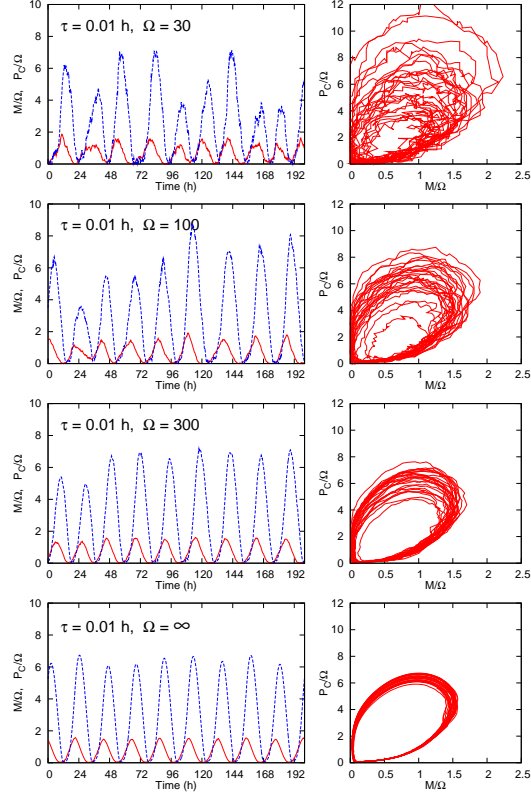


Figure 2: Oscillatory behaviors of the concentrations of mRNA and the cytosolic protein for various values of  $\Omega$  with  $\tau = 0.01$  h. The plots in the left column show the time variation and those in the right column are the projections of the trajectories in the  $M/\Omega - P_C/\Omega$  plane. We employ the same reaction parameters with those in Gonze et al.:  $n = 4$ ,  $v_s = 0.5$  nM h<sup>-1</sup>,  $K_I = 2.0$  nM,  $v_m = 0.3$  nM h<sup>-1</sup>,  $K_m = 0.2$  nM,  $k_s = 2.0$  h<sup>-1</sup>,  $v_d = 1.5$  nM h<sup>-1</sup>,  $K_d = 0.1$  nM,  $k_1 = 0.2$  h<sup>-1</sup>,  $k_2 = 0.2$  h<sup>-1</sup>.

no.	reaction	transition rate
a	$\begin{matrix} G = 1 \\ P_N \end{matrix} \longrightarrow \begin{matrix} G = 0 \\ P_N - n \end{matrix}$	$\frac{1}{\tau} \left( \frac{P_N}{K_I \Omega} \right)^n G$
b	$\begin{matrix} G = 0 \\ P_N \end{matrix} \longrightarrow \begin{matrix} G = 1 \\ P_N + n \end{matrix}$	$\frac{1}{\tau} (1 - G)$
1	$M \longrightarrow M + 1$	$v_s \Omega G$
2	$M \longrightarrow M - 1$	$v_m \Omega \frac{M/\Omega}{K_m + M/\Omega}$
3	$P_C \longrightarrow P_C + 1$	$k_s M$
4	$P_C \longrightarrow P_C - 1$	$v_d \Omega \frac{P_C/\Omega}{K_d + P_C/\Omega}$
5	$\begin{matrix} P_C \\ P_N \end{matrix} \longrightarrow \begin{matrix} P_C - 1 \\ P_N + 1 \end{matrix}$	$k_1 P_C$
6	$\begin{matrix} P_C \\ P_N \end{matrix} \longrightarrow \begin{matrix} P_C + 1 \\ P_N - 1 \end{matrix}$	$k_2 P_N$

Table 1: Reaction table for a simplified circadian system.

In order to see the effect of gene stochasticity, we examine the case for various values of  $\tau$  in the large  $\Omega$  limit (Fig.3). The fluctuation decreases on decreasing  $\tau$  as in the case of increasing  $\Omega$ . The trajectories are smoother in comparison with the previous case because the stochasticity is limited to the gene activity. One can see that the fluctuation is evident even in the case  $\tau = 0.01\text{h}$ , where the ratio  $\tau$  to the period ( $\sim 24\text{h}$ ) is as small as  $0.5 \times 10^{-3}$ .

Figure 4 shows the period (i.e. the peak-to-peak interval) distributions and the peak value distributions of  $[P_C]$  for  $\tau = 1.0, 0.1$ , and  $0.01\text{ h}$ . The averages and the standard deviations for the distributions are tabulated in Table 2. Both of the distributions becomes narrower for the smaller value of  $\tau$  approximately as  $\sqrt{\tau}$ , but the standard deviation of the period distribution



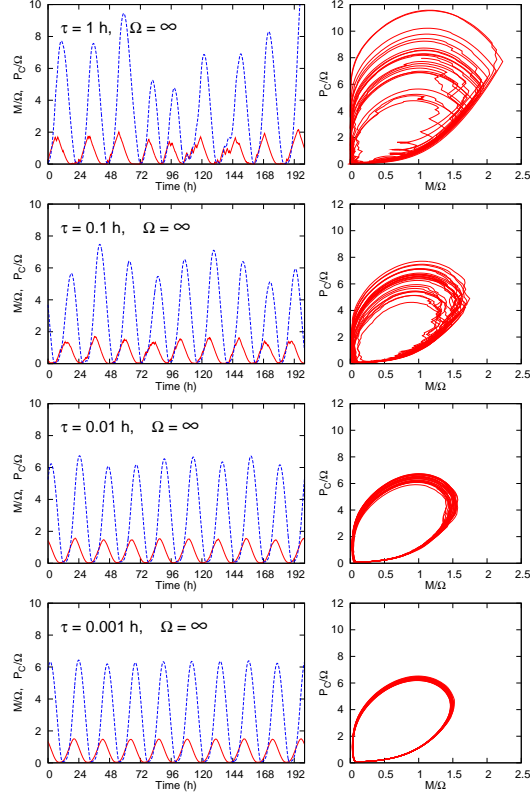


Figure 3: Oscillatory behaviors of the concentrations of mRNA and the cytosolic protein for various values of  $\tau$  with  $\Omega = \infty$ . The parameters are the same with those in Fig.2.

is still about a half hour even for the case of  $\tau = 0.01$  h. It should be noted that the ratios of the standard deviation to the average for the peak value distributions are about twice as large as those for the period distributions. More systematic data are presented in Supplementary material to show the  $\sqrt{\tau}$  scaling and the ratio of the two distribution widths.

The time correlation function  $C(t)$  of the nuclear protein concentration  $[P_N(t)]$  is defined as

$$C(t) = \frac{1}{T} \int_0^T \Delta[P_N(t' + t)] \Delta[P_N(t')] dt', \quad (17)$$

where  $\Delta[P_N(t)]$  represents the deviation from the average,

$$\Delta[P_N(t)] \equiv [P_N(t)] - \frac{1}{T} \int_0^T [P_N(t')] dt' \quad (18)$$

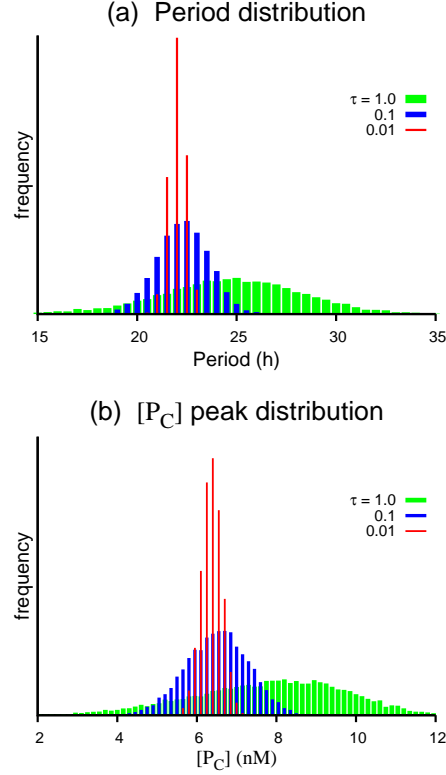


Figure 4: Distributions for (a) the period (i.e. peak-to-peak interval) and (b) the peak value of the cytosolic protein variation for  $\tau = 1.0, 0.1, 0.01$  h with  $\Omega = \infty$ . The averages and the standard deviations are tabulated in Table 2, from which one can see that the width of the distribution scales roughly as  $\sqrt{\tau}$ .

$\tau$ (h)		1	0.1	0.01
period (h)	av.	24.6	22.3	22.0
	std.	3.92	1.32	0.45
	std./av.	0.16	0.059	0.020
peak value (nM)	av.	7.82	6.50	6.37
	std.	1.87	0.79	0.26
	std./av.	0.24	0.12	0.041

Table 2: The averages and the standard deviations for the period and the peak value distributions for the cytosolic protein concentration shown in Fig.4. The ratios of the standard deviation to the average for the peak value distributions are about twice as large as those for the period distributions.

with  $T$  being the time length of the whole simulation. In Fig.5, the correlation functions are plotted and fitted to the form of damped oscillation

$$A \cos(\omega_0 t + \theta_0) e^{-t/\tau_{\text{corr}}} \quad (19)$$

to estimate the correlation time  $\tau_{\text{corr}}$ . Figure 6 shows the  $\tau$  dependence of the correlation time  $\tau_{\text{corr}}$  in the logarithmic scale. It shows the scaling

$$\tau_{\text{corr}} \sim \frac{1}{\tau} \quad (20)$$

in the small  $\tau$  regime, and the longer correlation time in the  $n = 4$  case than in the  $n = 1$  case. One may notice that the correlation time for  $\tau = 0.01$  h is quite long, i.e.  $\tau_{\text{corr}} \approx 2000$  h for the  $n = 4$  case, even though the period fluctuations are substantial as can be seen in Fig. 4 (See Appendix).

The scaling of  $\tau_{\text{corr}}$  given by Eq.(20) can be understood as the phase diffusion when the standard deviations of the period distributions scales as  $\sqrt{\tau}$  as shown in Fig. 4.

## Discussion

We have examined the effects of molecular fluctuations in a biological system on a simplified model of a circadian rhythm system, where there are two types of fluctuation sources: (i) small numbers of molecules involved and (ii) finite time scale of the gene regulation. The first effect has been studied by Gonze et al. (2002a), assuming that the gene regulation time scale is infinitesimal. In the present work, we focus on the second effect, i.e., in the case where the molecular numbers are large enough that the fluctuation due to the first effect is negligible.

We have developed a method to study this effect systematically by introducing a new parameter  $\tau$  to scale the gene regulation times. We set  $\tau$  to be the unbinding time of the transcription factor, keeping the ratio of the binding to the unbinding rate constant. We performed numerical simulations for various values of  $\tau$  without an external entrainment of the 24-hour period. As  $\tau$  decreases, the oscillation appears more deterministic; the width of the distributions for the oscillation periods and the peak values scales with  $\sqrt{\tau}$  and the correlation time for the correlations function scales with  $1/\tau$ . We have found that the system is very sensitive to such fluctuation, and demonstrated that the oscillation period fluctuates by about 30 min even for very

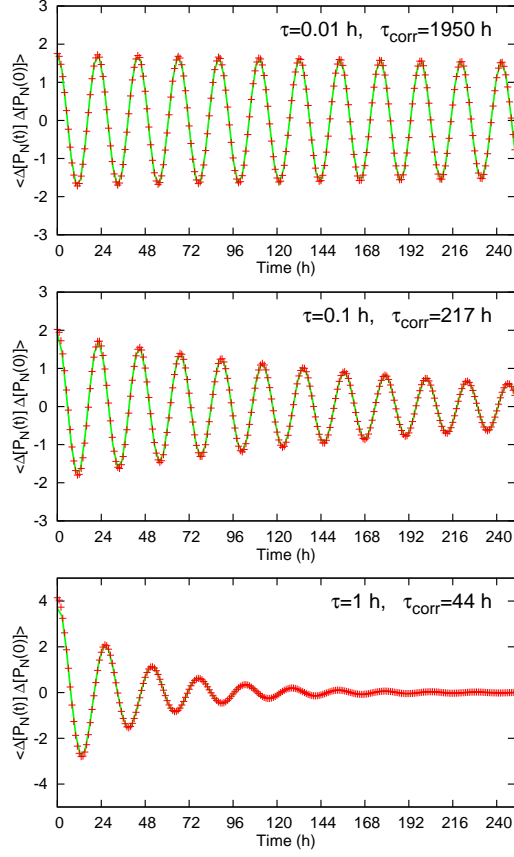


Figure 5: The time correlation functions for the nuclear protein concentration for the regulation time  $\tau=0.01, 0.1, 1$  h with  $\Omega = \infty$ . The (green) lines shows the fitting curves of the form  $A \cos(\omega_0 t + \theta_0) e^{-t/\tau_{\text{corr}}}$ . The fitted values of  $\tau_{\text{corr}}$  are shown on the plots. The other parameters are the same with those in Fig.2.

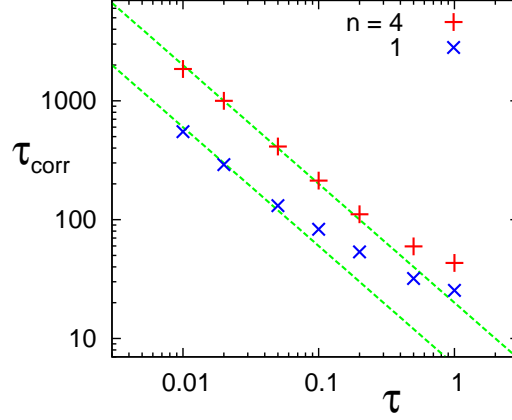


Figure 6: The regulation time  $\tau$  dependences of the correlation time  $\tau_{\text{corr}}$  for  $n = 1$  and 4 with  $\Omega = \infty$  in the logarithmic scale. The (green) dashed lines shows the fitted lines proportional to  $\tau^{-1}$ .

small  $\tau = 0.01 \text{ h} \approx 30 \text{ s}$  in comparison with its period around 22 h. For the present parameter set, the nuclear protein concentration  $P_N/\Omega$  oscillates in the range  $0 \sim 5 \text{ nM}$ , therefore, the value of  $\tau \approx 30 \text{ s}$  for the unbinding time gives about 1 s for the binding time. These estimates may be tested with experimental data.

The 30 minutes period fluctuation is large for a circadian system. This sensitivity to the fluctuation in the gene regulation is an interesting feature of the present simplified model. Multiple feedback loops with several phosphorylation steps found in actual biological systems may be designed in such a way as to reduce this sensitivity (Ueda et al., 2001; Kusakina and Dodd, 2012). This can be studied by extending the present method.

The correlation function for the protein oscillation fits to the damped sinusoidal function very well, and the estimated correlation time  $\tau_{\text{corr}}$  scales as  $1/\tau$  in the small  $\tau$  regime. Such decay in the correlation function is caused by the phase diffusion due to fluctuations. We estimate  $\tau_{\text{corr}}$  for the Hill coefficient for the gene regulation  $n=4$  and 1, and found that  $\tau_{\text{corr}}$ 's for  $n=4$  are about 5 times larger than those for  $n=1$ ; the fluctuation effect is suppressed by the larger value of the Hill coefficient by the cooperativity effects as in the case of Gonze et al. (2002b,a).

It is also interesting to find that the relative fluctuations for the peak values are twice as large as those for the periods, namely, the period is more stable than the amplitude. This may be a reason why the correlation time

is quite long in spite of apparent fluctuations in the oscillation.

In the present work, we study only the case where the copy number of the gene is 1. However, there are typically a couple of genes in a cell. In the case of multiple genes in a cell, the fluctuation in each gene cancels each other, therefore overall fluctuation will be reduced. We confirmed by simulation that the fluctuation for a two-gene system with  $\tau$  is almost the same as that for a single gene system with  $\tau/2$ . This is because the fluctuation cancellation by two genes should be comparable with that by one gene that switches twice as fast. Simulation data are presented in Supplementary material.

The fluctuation indicated by our simulations may be compared with previous experimental observations. Although circadian clocks are very accurate as a system, large fluctuations have been observed in the oscillation of individual cells of fibroblasts (Nagoshi et al., 2004) and cyanobacteria (Mihalcescu et al., 2004) when they oscillate independently. For both cases, it is reported that the fluctuations are much larger for the amplitude than those for the period. In the latter case (Mihalcescu et al., 2004), the correlation time is estimated as long as  $166 \pm 100$  days in spite of apparent large fluctuations in the amplitude. Such a long correlation time corresponds to our case of the gene regulation time scale  $\tau = 0.01$  h, which gives the  $\tau_{\text{corr}} \approx 1950$  h.

Very little fluctuation is usually observed in circadian systems; fluctuation in the period is typically less than 10 minutes (Amdaoud et al., 2007), which is even smaller than the fluctuation of 30 minutes that we obtained for the case  $\tau = 0.01$  h. There are some possible mechanisms to suppress molecular fluctuations. (i) Cooperativity among cells: The present system models a single cell behavior, but for the case of multicellular organisms, the cooperativity among cells may exist and that should reduce the fluctuation in each cell. Actually, variability in each cell is much larger than that of a whole system in the case of multicellular organisms (Liu et al., 1997; Yamaguchi et al., 2003; Herzog et al., 2004; Carr and Whitmore, 2005). (ii) Multiple feedback loops: Our model is a simplified core model for a circadian system and consists of a single negative feedback loop. However, it has been known that circadian systems typically consist of multiple feedback loops (Zeng et al., 1996; Leloup and Goldbeter, 1998; Glossop et al., 1999; Blau, 2001), which could be designed in the way to compensate the fluctuations in one loop by the other. (iii) Chemical oscillation without gene control: In the case of cyanobacteria, it has been proposed that the circadian system consists of proteins only and does not involve a gene expression (Tomita et al., 2005).

In such a system, the fluctuation discussed in this work does not exist.

Other than circadian systems, there are some oscillations observed in biology such as Hes1 oscillation during somite segmentation (Hirata et al., 2002), p53 oscillation after DNA damage by gamma irradiation (Geva-Zatorsky et al., 2006), or oscillations in artificially constructed systems (Elowitz and Leibler, 2000; Atkinson et al., 2003; Tigges et al., 2009). In these systems, the fluctuations are much more profound than circadian systems, and part of the fluctuations should come from the gene regulatory processes, for which the present analysis is applicable.

In summary, we have developed a theoretical tool to study the molecular fluctuation due to the finite transcription regulation time, and have demonstrated that a symplified core model of circadian system is sensitive to such fluctuation. Our method can be extended to study a more realistic system and can be utilized to clarify biological significance of a detailed design of circadian system in terms of stability against the molecular fluctuation.

## Appendix: Phase diffusion and correlation time

In the appendix, we show that the correlation time  $\tau_{\text{corr}}$  in the correlation function is proportional to  $1/\tau$  when the period distribution has the width proportional to  $\sqrt{\tau}$ .

Suppose the correlation function  $C(t)$  is written as

$$C(t) = A \int_{-\infty}^{\infty} \cos(\omega_0 t + \theta) P(\theta, t) d\theta \quad (\text{A.1})$$

in terms of the average over the phase difference  $\theta$  by the distribution function  $P(\theta, t)$  at time  $t$ . Here,  $\omega_0$  is the average angular velocity given by  $\omega_0 = 2\pi/T_0$  in terms of the average period  $T_0$ . Now, we assume that the phase distribution can be approximated by the Gaussian distribution with the standard deviation  $\sigma_\theta(t)$ ,

$$P(\theta, t) \approx \frac{1}{\sqrt{2\pi\sigma_\theta^2(t)}} \exp \left[ -\frac{\theta^2}{2\sigma_\theta^2(t)} \right], \quad (\text{A.2})$$

then, Eq.(A.1) may be estimated as

$$C(t) \approx A \cos(\omega_0 t) \exp \left[ -\frac{1}{2}\sigma_\theta^2(t) \right]. \quad (\text{A.3})$$

Now, we estimate the phase distribution  $P(\theta, t)$  as follows. The phase  $\theta$  at the time  $t = nT_0$  may be expressed as the sum of  $n$  phases accumulated by the time:

$$\theta = \sum_{i=1}^n 2\pi \left( \frac{1}{T_i} - \frac{1}{T_0} \right) T_0 \approx - \sum_{i=1}^n 2\pi \frac{\Delta T_i}{T_0}, \quad (\text{A.4})$$

where  $T_i$  is the  $i$ 'th period (i.e. peak-to-peak interval) with  $T_i = T_0 + \Delta T_i$ , and we have assumed  $\Delta T_i \ll T_0$  in the last approximation.

If there is no correlation among  $\Delta T_i$ , then  $\sigma_\theta(t)$  is given by

$$\sigma_\theta^2(t) \approx \left( 2\pi \frac{\sigma_T}{T_0} \right)^2 \frac{t}{T_0}, \quad (\text{A.5})$$

where  $\sigma_T$  is the standard deviation of the period  $T$ . We have replaced  $n$  by  $t/T_0$ .

With Eq.(A.3), this gives the form of Eq.(19) with  $\theta_0 = 0$  and

$$\tau_{\text{corr}} = \left( \frac{1}{2\pi} \frac{T_0}{\sigma_T} \right)^2 2T_0, \quad (\text{A.6})$$

thus if  $\sigma_T \propto \sqrt{\tau}$ , we obtain  $\tau_{\text{corr}} \propto 1/\tau$ .

**Acknowledgments:** The authors would like to acknowledge Dr. Hiroshi Ito for informative discussions.

## References

- Amdaoud, M., Vallade, M., Weiss-Schaber, C., Mihalcescu, I., 2007. Cyanobacterial clock, a stable phase oscillator with negligible intercellular coupling. PNAS 104, 7051–7056.
- Atkinson, M.R., Savageau, M.A., Myers, J.T., Ninfa, A.J., 2003. Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in escherichia coli. Cell 113, 597–607.
- Barkai, N., Leibler, S., 2000. Biological rhythms: Circadian clocks limited by noise. Nature 403, 267–268.
- Blau, J., 2001. The drosophila circadian clock: what we know and what we don't know. Semin. Cell Dev. Biol. 12, 287–293.



- Carr, A.J.F., Whitmore, D., 2005. Imaging of single light-responsive clock cells reveals fluctuating free-running periods. *Nature Cell Biology* 7, 319–321.
- Dunlap, J.C., 1999. Molecular bases for circadian clocks. *Cell* 96, 271–290.
- Elowitz, M.B., Leibler, S., 2000. A synthetic oscillatory network of transcriptional regulators. *Nature* 403, 335–338.
- Geva-Zatorsky, N., Rosenfeld, N., Itzkovitz, S., Milošević, R., Sigal, A., Dekel, E., Yarnitzky, T., Liron, Y., Polak, P., Lahav, G., Alon, U., 2006. Oscillations and variability in the p53 system. *Molecular Systems Biology* 2.
- Gillespie, D.T., 1976. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* 22, 403–434.
- Gillespie, D.T., 1977. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* 81, 2340–2361.
- Glossop, N.R.J., Lyons, L.C., Hardin, P.E., 1999. Interlocked feedback loops within the drosophila circadian oscillator. *Science* 286, 766–768.
- Goldbeter, A., 1995. A model for circadian oscillations in the drosophila period protein (per). *Proc. R. Soc. Lond. B* 261, 319–324.
- Goldbeter, A., 1996. *Biochemical Oscillations and Cellular Rhythms: The Molecular Bases of Periodic and Chaotic Behaviour*. Cambridge University Press, Cambridge, UK.
- Gonze, D., Halloy, J., Gaspard, P., 2002a. Biochemical clocks and molecular noise: Theoretical study of robustness factors. *J. Chem. Phys.* 116, 10997–11010.
- Gonze, D., Halloy, J., Goldbeter, A., 2002b. Robustness of circadian rhythms with respect to molecular noise. *PNAS* 99, 673–678.
- Herzog, E.D., Aton, S.J., Numano, R., Sakaki, Y., Tei, H., 2004. Temporal precision in the mammalian circadian system: A reliable clock from less reliable neurons. *J. Biol. Rhythms* 19, 35–46.

- Hirata, H., Yoshiura, S., Ohtsuka, T., Bessho, Y., Harada, T., Yoshikawa, K., Kageyama, R., 2002. Oscillatory expression of the bhlh factor *hes1* regulated by a negative feedback loop. *Science* 298, 840–843.
- Kusakina, J., Dodd, A.N., 2012. Phosphorylation in the plant circadian system. *Trends in Plant Science*, to be published.
- Leloup, J.C., Goldbeter, A., 1998. A model for circadian rhythms in *drosophila* incorporating the formation of a complex between the *per* and *tim* proteins. *J. Biol. Rhythms* 13, 70–87.
- Leloup, J.C., Gonze, D., Goldbeter, A., 1999. Limit cycle models for circadian rhythms based on transcriptional regulation in *drosophila* and *neurospora*. *J. Biol. Rhythms*. 14, 433–448.
- Liu, C., Weaver, D.R., Strogatz, S.H., Reppert, S.M., 1997. Cellular construction of a circadian clock: Period determination in the suprachiasmatic nuclei. *Cell* 91, 855–860.
- Mihalcescu, I., Hsing, W., Leibler, S., 2004. Resilient circadian oscillator revealed in individual cyanobacteria. *Nature* 430, 81–85.
- Nagoshi, E., Saini, C., Bauer, C., Laroche, T., Naef, F., Schibler, U., 2004. Circadian gene expression in individual fibroblasts: Cell-autonomous and self-sustained oscillators pass time to daughter cells. *Cell* 119, 693–705.
- Nicolis, G., 1972. Fluctuations around nonequilibrium states in open nonlinear systems. *J. Stat. Phys.* 6, 195–222.
- Tigges, M., Marquez-Lago, T.T., Stelling, J., Fussenegger, M., 2009. A tunable synthetic mammalian oscillator. *Nature* 457, 309–312.
- Tomita, J., Nakajima, M., Kondo, T., Iwasaki, H., 2005. No transcription-translation feedback in circadian rhythm of *kaic* phosphorylation. *Science* 307, 251–254.
- Ueda, H.R., Hagiwara, M., Kitano, H., 2001. Robust oscillations within the interlocked feedback model of *drosophila* circadian thythm. *J. theor. Biol.* 210, 401–406.

- Yamaguchi, S., Isejima, H., Matsuo, T., Okura, R., Yagita, K., Kobayashi, M., Okamura, H., 2003. Synchronization of cellular clocks in the suprachiasmatic nucleus. *Science* 302, 1408–1412.
- Young, M.W., 2000. Life's 24-hour clock: molecular control of circadian rhythms in animal cells. *Trends Biochem. Sci.* 25, 601–606.
- Zeng, H., Qian, Z., Myers, M.P., Rosbash, M., 1996. A light-entrainment mechanism for the drosophila circadian clock. *Nature* 380, 129–135.

Supplementary material to  
“Transcription fluctuation effects on biochemical oscillations”  
by Ryota Nishino, Takahiro Sakaue, and Hiizu Nakanishi

In this supplementary material, we present data to demonstrate the  $\sqrt{\tau}$ -scaling of the distribution width and the multiple gene effects more systematically.

### S.1. $\sqrt{\tau}$ -scaling of the distribution width

In the text, we show only three sets of data for the distributions of the period and the peak value of  $[P_C]$  in order to demonstrate the  $\sqrt{\tau}$  scaling of the distribution width. We examined this scaling more systematically.

In Fig.S.1, the ratios of the standard deviation to the average are plotted against  $\tau$  in the logarithmic scale. The dashed lines denote the fitting lines with the slope 0.5, which shows the  $\sqrt{\tau}$  scaling of the distribution width of the oscillation parameters in the small  $\tau$  region.

This can be understood naturally; The on/off frequency of the gene is proportional to  $1/\tau$ , thus the fluctuation in the total on/off-time length scales with  $\sqrt{\tau}$ , from which we expect that the distribution width of the oscillation parameters scales with  $\sqrt{\tau}$  because the response in fluctuation should be linear in the small input fluctuation limit.

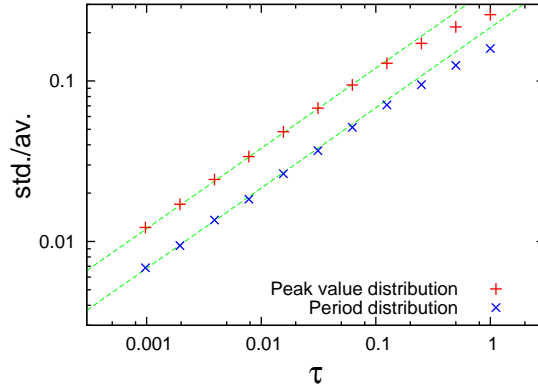


Figure S.1: Scaling of the distribution width by  $\tau$ . The ratio of the standard deviation to the average is plotted against  $\tau$  in the logarithmic scale for the peak value of  $[P_C]$  and the period distributions. The dashed lines are the lines with the slope 0.5 fitted to each data set.

## S.2. Multiple gene effects

In the paper, we only consider the case where the copy number of the gene is one. Here, we present the analysis for the case where the copy number of the gene is  $G_0$ . The transition rate for each reaction is given by Table S.1, where  $G$  is the number of active genes. Note that the transition rate for the reaction 1 is scaled by  $1/G_0$  in order to give the same rate equation as before in the limit of  $\tau \rightarrow 0$  and  $\Omega \rightarrow \infty$ .

In the  $\tau \rightarrow 0$  limit, the time dependent average of  $G$  denoted by  $G_{\text{Av}}(t)$  is given by

$$G_{\text{Av}}(t) = G_0 \frac{1}{1 + ([P_N]/K_I)^n}, \quad (\text{S.1})$$

which gives the same rate equations for the concentrations if the transition rate for the reaction 1 is scales by  $1/G_0$ .

To see the copy number effects, we performed Monte Carlo simulations for  $G_0 = 1$  and 2 with some values of  $\tau$  in the  $\Omega \rightarrow \infty$  limit. Fig. S.2 shows the distributions for the period and the peak values of  $[P_C]$ . One can see the distribution for  $G_0 = 2$  with  $\tau = 0.1$  h agrees with that for  $G_0 = 1$  with  $\tau = 0.05$  h quite well for both of the distributions.

Actually, this can be understood in a simple way; For the copy number  $G_0 = 2$ , the fluctuations in the activity of the two genes cancel each other. This cancellation should be comparable to the fluctuation cancellation in the system with  $G_0 = 1$  and the half time scale because the gene activity switches between on and off twice as fast.

no.	reaction	transition rate
a	$\begin{matrix} G \\ P_N \end{matrix} \longrightarrow \begin{matrix} G-1 \\ P_N-n \end{matrix}$	$\frac{1}{\tau} \left( \frac{P_N}{K_I \Omega} \right)^n G$
b	$\begin{matrix} G \\ P_N \end{matrix} \longrightarrow \begin{matrix} G+1 \\ P_N+n \end{matrix}$	$\frac{1}{\tau} (G_0 - G)$
1	$M \longrightarrow M+1$	$v_s \Omega \frac{G}{G_0}$
2	$M \longrightarrow M-1$	$v_m \Omega \frac{M/\Omega}{K_m + M/\Omega}$
3	$P_C \longrightarrow P_C+1$	$k_s M$
4	$P_C \longrightarrow P_C-1$	$v_d \Omega \frac{P_C/\Omega}{K_d + P_C/\Omega}$
5	$\begin{matrix} P_C \\ P_N \end{matrix} \longrightarrow \begin{matrix} P_C-1 \\ P_N+1 \end{matrix}$	$k_1 P_C$
6	$\begin{matrix} P_C \\ P_N \end{matrix} \longrightarrow \begin{matrix} P_C+1 \\ P_N-1 \end{matrix}$	$k_2 P_N$

Table S.1: Reaction table for a simplified circadian system in the case where the copy number of the gene is  $G_0$ .

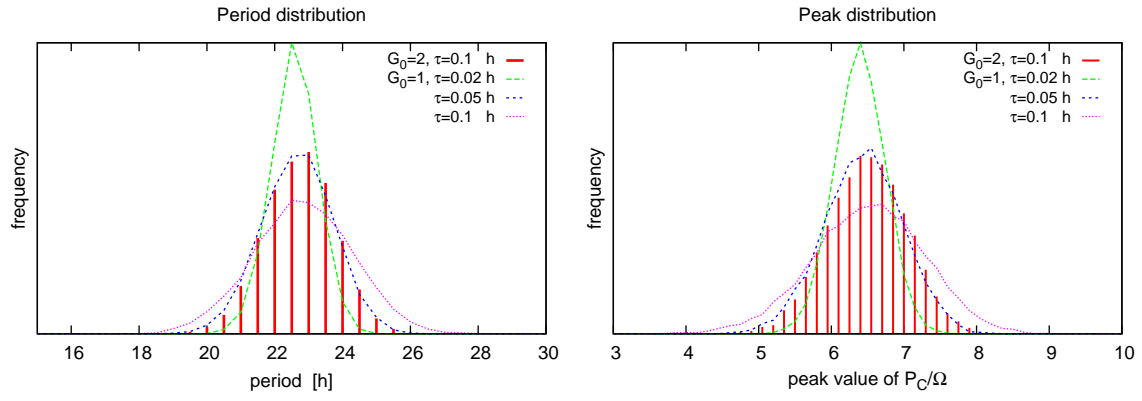


Figure S.2: Period and  $[P_C]$  peak value distributions for  $G_0 = 1$  and 2 with various values of  $\tau$  in the  $\Omega \rightarrow \infty$  limit. The other parameters are the same with those in the text:  $n = 4$ ,  $v_s = 0.5 \text{ nM h}^{-1}$ ,  $K_I = 2.0 \text{ nM}$ ,  $v_m = 0.3 \text{ nM h}^{-1}$ ,  $K_m = 0.2 \text{ nM}$ ,  $k_s = 2.0 \text{ h}^{-1}$ ,  $v_d = 1.5 \text{ nM h}^{-1}$ ,  $K_d = 0.1 \text{ nM}$ ,  $k_1 = 0.2 \text{ h}^{-1}$ ,  $k_2 = 0.2 \text{ h}^{-1}$ .